

Embedding in q -ary 1-perfect codes and partitions*

D. S. Krotov[†], E. V. Sotnikova[‡]

December 16, 2014

Abstract

We prove that every 1-error-correcting code over a finite field can be embedded in a 1-perfect code of some larger length. Embedding in this context means that the original code is a subcode of the resulting 1-perfect code and can be obtained from it by repeated shortening. Further, we generalize the results to partitions: every partition of the Hamming space into 1-error-correcting codes can be embedded in a partition of a space of some larger dimension into 1-perfect codes. For the partitions, the embedding length is close to the theoretical bound for the general case and optimal for the binary case.

1 Introduction

The goal of the current work is to show that any (in general, nonlinear) code that can correct at least one error is a subcode of a 1-perfect code of some larger length. Moreover, we prove a similar result for the partitions into 1-error-correcting codes. In [1], the possibility to embed into 1-perfect code was proven for binary codes. In [4], the ternary case was solved; for the q -ary case with $q > 3$, there are similar embedding results in [4], but with the restriction that the embedded code is required to be at least 2-error-correcting (this restriction is very strict as almost all 1-error-correcting codes are not 2-error-correcting). The reason of such restriction is that the

*The work was financed by the Russian Science Foundation (grant No 14-11-00555)

[†]Sobolev Institute of Mathematics, Novosibirsk, Russia; Novosibirsk State University, Novosibirsk, Russia. E-mail: krotov@math.nsc.ru

[‡]Sobolev Institute of Mathematics, Novosibirsk, Russia. E-mail: lucernavesper@gmail.com

method suggested in [1] does not work in general case: the components that should be switched in the linear 1-perfect code to build the required subcode can intersect in the case $q > 3$ (see Remark 1). To avoid this problem, we suggest a modification of the method.

We will follow the convenient notation and line of reasoning from [1] with three main differences. At first, the key definition of a linear i -component (in our notation, we will write a Greek letter instead of traditional i) is now given in a usual form [3], while the required property is declared in Lemma 2 (the definition based on this property would look complicate in the q -ary case). At second, the formulation of the crucial proposition, which is essentially the main and largest part of the proof of the main theorem, is different from the crucial lemma in the binary case (as was noted above, the last one does not work in the general case, see also Remark 1). At third, we add the theorem about embedding partitions, which is new for all q , including $q = 2$.

2 Notation and definitions

Over the article, we will use the following notation.

- F denotes the Galois field $\text{GF}(q)$ of order q .
- F^m is the set of m -tuples over F , considered as a vector space over F . The elements of F^m are denoted by Greek letters.
- \mathfrak{A} consists of m -tuples from F^m with the first nonzero element equal to 1.
- The intersection of \mathfrak{A} with a 2-dimensional subspace of F^m will be referred to as a *line*. The cardinality of every line is $q + 1$. The set of lines together with the set of the *points* \mathfrak{A} form an incidence structure, known as the *projective geometry* $\text{PG}(m - 1, q)$.
- The intersection of \mathfrak{A} with a 3-dimensional subspace F^m will be referred to as a *plane*.
- $n \stackrel{\text{df}}{=} \frac{q^m - 1}{q - 1}$.
- $\Pi = \{\pi^{(1)}, \dots, \pi^{(m)}\} = \{(1 \ 0 \ \dots \ 0), \dots, (0 \ \dots \ 0 \ 1)\}$ is the natural basis in F^m .
- The elements of F^n will be denoted by overlined letters with the coordinates indexed by the elements of \mathfrak{A} . We assume that the first m

coordinates have the indexes $\pi^{(1)}, \dots, \pi^{(m)}$, while the other $n - m$ coordinates are ordered in some arbitrary fixed way.

- $\{\bar{e}^{(\delta)}\}_{\delta \in \mathfrak{A}}$ is the natural basis in F^n , herewith $\bar{e}^{(\pi^{(\delta)})} = (\pi^{(\delta)}, 0^{n-m})$.
- For any $\alpha = (\alpha_1, \dots, \alpha_m) \in F^m$, we define $\bar{\alpha} \stackrel{\text{df}}{=} (\alpha, 0^{n-m}) \in F^n$; moreover $\bar{\alpha} = \sum_{i=1}^m \alpha_i \bar{e}^{(\pi^{(i)})}$.
- The *Hamming distance* $d(x, y)$ is the number of positions in which vectors x, y from the same space differ.
- The *neighborhood* $\Omega(M)$ of a set $M \subset F^n$ is the set of the vectors at distance at most 1 from M .
- A set $C \subset F^n$ is called a *1-code* if the neighborhoods of the codewords are disjoint.
- A 1-code $P \subset F^n$ is called a *1-perfect code* if $\Omega(P) = F^n$.
- The *Hamming code* \mathcal{H}_m of length n is defined as the set of vectors $\bar{c} \in F^n$ satisfying the following equation:

$$\sum_{\alpha \in \mathfrak{A}} c_\alpha \alpha = 0^m. \quad (1)$$

- $\text{supp}(c) = \{\delta \in \mathfrak{A} \mid c_\delta \neq 0\}$.
- $T \stackrel{\text{df}}{=} \{c \in \mathcal{H}_m \mid |\text{supp}(c)| = 3\}$. The elements of T are called *triples*.
- $T_\delta \stackrel{\text{df}}{=} \{c \in T \mid c_\delta = 1\}$.
- The *linear δ -component* R_δ is defined as the linear span $\langle T_\delta \rangle$. By an *δ -component of the Hamming code*, we will mean any coset of the linear δ -component that is a subset of the Hamming code.

3 Preliminaries

Lemma 1. *For any $\bar{z} \in F^n$ it holds that $\Omega(R_\delta + \bar{z}) = \Omega(R_\delta + \bar{z} + \mu \bar{e}^{(\delta)})$ for all $\mu \in F$.*

Proof. Without loss of generality it is enough to prove the statement for $\bar{z} = 0^n$. It is shown in [3] that $(\mathcal{H}_m \setminus R_\delta) \cup (R_\delta + \mu \bar{e}^{(\delta)})$ is a 1-perfect code for all $\mu \in F$. From the definition of the 1-perfect code it follows that the neighborhoods of the sets R_δ and $R_\delta + \mu \bar{e}^{(\delta)}$ are equal. So the statement of Lemma is true. \square

Lemma 2. *Let $\delta \in \mathfrak{A}$. Every word \bar{c} from R_δ satisfies the relation*

$$\sum_{\alpha \in \mathfrak{L}} c_\alpha l(\alpha) = 0^m \quad (2)$$

for all linear functions l from F^m to F such that $l(\delta) = 0$ and all lines \mathfrak{L} containing δ .

Proof. Since R_δ is a subset of the Hamming code, each of its elements \bar{c} satisfies (1). Then

$$\sum_{\alpha \in \mathfrak{A}} c_\alpha l(\alpha) = 0^m \quad (3)$$

holds for any linear function l . Now assume $l(\delta) = 0$ and consider a line \mathfrak{L} containing δ . Then, the support of every triple from T_δ either is included in \mathfrak{L} or intersect with \mathfrak{L} in only one element δ . In the last case, (2) is trivial; in the former case, it trivially follows from (3). Since the required relation holds for every element of T_δ , we see from linearity that it holds for the linear span of T_δ , i.e., for R_δ . \square

Lemma 3. *Let $\delta, \kappa \in \mathfrak{A}$. Every element \bar{c} of $\langle R_\delta, R_\kappa \rangle$ satisfies the relation*

$$\sum_{\alpha \in \mathfrak{P}} c_\alpha l(\alpha) = 0^m \quad (4)$$

for all linear functions l from F^m to F such that $l(\delta) = l(\kappa) = 0$ and all planes \mathfrak{P} containing δ and κ .

Proof. First consider the case $\bar{c} \in R_\delta$. Summarizing (2) over the all lines containing δ and included in \mathfrak{P} , we get (4). So, the elements of R_δ and, similarly, the elements of R_κ satisfy (4). By the linearity, the elements of $\langle R_\delta, R_\kappa \rangle$ do. \square

4 Embedding in a 1-perfect code

Proposition 1. *Assume that δ and κ from F^m both start with 1 and the distance between them is at least 3. Then the δ -component $R_\delta + \bar{\delta} - \bar{e}^\delta$ and the κ -component $R_\kappa + \bar{\kappa} - \bar{e}^\kappa$ are disjoint.*

Proof. Consider the vector difference $\bar{c} = (\bar{\delta} - \bar{e}^{(\delta)}) - (\bar{\kappa} - \bar{e}^{(\kappa)})$. It is sufficient to show that $\bar{c} \notin \langle R_\delta, R_\kappa \rangle$. We will show that \bar{c} does not satisfy (4). Note that the first element of \bar{c} is 0, and $c_{\pi(i)} \neq 0$ if and only if $\delta_i \neq \kappa_i$. Among the other coordinates (not from Π), \bar{c} has exactly two nonzero positions, δ

and κ . Now consider some i such that $c_{\pi^{(i)}} \neq 0$. Note that $\pi^{(i)}$, δ and κ are linearly independent (indeed, a nontrivial linear combination of δ and κ is either nonzero in the first position or a multiple of $\delta - \kappa$, which has at least three nonzeros and thus cannot coincide with $\pi^{(i)}$); hence there is a unique plane P containing $\pi^{(i)}$, δ and κ .

Now we state that $\pi^{(i)}$, δ and κ are the only points of P in which \bar{c} is not equal to zero. Indeed, assume that $\beta = h\pi^{(i)} + a\delta + b\kappa \in \mathfrak{A}$. If $a + b \neq 0$ then $\beta_1 \neq 0$ and thus either $\beta \in \{\delta, \kappa\}$ or $c_\beta = 0$ holds. If $a + b = 0$ then $a\delta + b\kappa = a(\delta - \kappa)$ and thus, by the hypothesis of the proposition, this combination has at list three nonzero positions. In this case, β has at least two nonzero positions, and thus does not belong to Π . Hence, $c_\beta = 0$.

Then we consider a linear function l such that $l(\delta) = l(\kappa) = 0 \neq l(\pi^{(i)})$ and see that (4) cannot hold as it has only one nonzero summand, $\alpha = \pi^{(i)}$. \square

Remark 1. The hypothesis that both δ and κ start with 1 is necessary in Proposition 1 for $q > 3$. For example let us consider $\delta = (1, 1, 1)$ and $\kappa = (t, t^2, t^2) = t(1, t, t) = t\gamma$, where t^2 is different from 1 and t (so, $q \geq 4$). Then the vectors $\bar{\delta}$ and $\bar{\kappa}$ are at distance 1 from the δ -component $R_\delta + \bar{\delta} - \bar{e}^\delta$ and the γ -component $R_\gamma + \bar{\kappa} - t\bar{e}^\gamma$ of the Hamming code. It is easy to see that the nonzero coordinates $\pi^{(1)}$, $\pi^{(2)}$, $\pi^{(3)}$, δ and γ of the difference $\bar{c} = (\bar{\delta} - \bar{e}^{(\delta)}) - (\bar{\kappa} - t\bar{e}^{(\gamma)})$ belong to the same plain. Hence, since this difference is from the Hamming code, we see that it satisfies (4). It is not difficult to conclude that the corresponding components intersect.

Theorem 1. *Let $C \subset F^{m-1}$ be a 1-code. Define $\dot{C} \stackrel{\text{df}}{=} \{(1, x) \mid x \in C\}$. Then the following set*

$$P(C) \stackrel{\text{df}}{=} \left(\mathcal{H}_m \setminus \bigcup_{\delta \in \dot{C}} (R_\delta + \bar{\delta} - \bar{e}^{(\delta)}) \right) \cup \left(\bigcup_{\delta \in \dot{C}} (R_\delta + \bar{\delta}) \right)$$

is a 1-perfect code in F^n , within

$$C = \{x \in F^{m-1} \mid (1, x, 0^{n-m}) \in P(C)\}. \quad (5)$$

Proof. It is clear that $\bar{\delta} - \bar{e}^{(\delta)} \in \mathcal{H}_m$ for all $\delta \in \mathfrak{A}$, which means $R_\delta + \bar{\delta} - \bar{e}^{(\delta)} \subset \mathcal{H}_m$ for all δ . According to Proposition 1 the sets $R_\delta + \bar{\delta} - \bar{e}^{(\delta)}$ are mutually disjoint for all $\delta \in \dot{C}$. As they are subsets of a 1-perfect code, their neighborhoods are also mutually disjoint. From Lemma 1 we see that $P(C)$ is a 1-perfect code.

To prove (5), we first note that $\bar{c} = (\alpha, 0^{n-m}) \in \mathcal{H}_m$ implies $\alpha = 0^m$, which follows from the definition of Hamming code. Finally, we need to show that

if for some $x \in F^{m-1}$ we have $(1, x, 0^{n-m}) \in R_\delta + \bar{\delta}$, then $(1, x) = \delta$. Indeed, if $(1, x, 0^{n-m}) \in R_\delta + \bar{\delta}$ then $(1, x, 0^{n-m}) - \bar{\delta} \in R_\delta \subset \mathcal{H}_m$, which only holds for $(1, x) = \delta$. This completes the proof. \square

5 Partitions

Theorem 2. *Let (C_1, \dots, C_k) be a partition of F^s into 1-codes. Then there is a partition $(P_1, \dots, P_{q^{s+1}})$ of F^n into 1-perfect codes of length $n = (q^{s+1} - 1)/(q - 1)$ such that for all $j = 1, \dots, k$,*

$$C_j = \{x \in F^s \mid (1, x, 0^{n-s-1}) \in P_j\}. \quad (6)$$

Proof. Put $m = s+1$. Let for all α from F^m , H_α be the coset of the Hamming code that contains $\bar{\alpha}$; so, $\{H_\alpha\}_{\alpha \in F^m}$ is a partition of F^n . Let us choose k distinct vectors y_1, \dots, y_k from F^s , and denote $\alpha_j = (0, y_j)$, $j = 1, \dots, k$. Using Theorem 1, we replace the code H_{α_j} by $P(C_j - y_j) + \bar{\alpha}_j$. Then, this code will be the j s element P_j of the constructed partition; readily, (6) is straightforward from (5). It remains to replace the other cosets of the Hamming code to get a partition. According to the definition of P_j , it intersects with the following cosets of the Hamming code: with H_{α_j} and, for every x from C_j , with $H_{(1,x)}$, which has a common component with P_j . Let O_x be obtained from $H_{(1,x)}$ by removing this component, and by including the corresponding component (the one that is not in P_j) of H_{α_j} . Now we see that the $|C_j| + 1$ codes H_{α_j} and O_x , $x \in C_j$, are mutually disjoint and

$$P_j \cup \bigcup_{x \in C_j} O_x = H_{\alpha_j} \cup \bigcup_{x \in C_j} H_{(1,x)}.$$

Then, the codes P_j , $j = 1, \dots, k$, together with the codes O_x , $x \in F^{m-1}$, and the codes H_α , where α does not start with 1 and is different from all α_j , $j = 1, \dots, k$, form a partition of F^n . As was noted above, (6) holds. \square

Note that, since the number k of codes in the original partition can be rather large, up to q^s , the length n for which it is possible to construct the embedding cannot be small too: the number $(q - 1)n + 1$ of perfect codes in the resulting partition cannot be smaller than q^s . So, $n \geq \frac{q^s - 1}{q - 1}$, and we see that our construction gives an embedding with “almost” minimal length $\frac{q^{s+1} - 1}{q - 1}$. Using the same approach and based on the results of [1] and [4], one can construct an embedding of minimal length for the cases $q = 2$ and $q = 3$, respectively.

Finally, we note that Theorem 2 is the most general known formulation that generalize Theorem 1 and, in particular, the result of [1] (putting aside

small increasing of the embedding length) and some results of [4]. As noted in [1], the classical results [5] and [2] about embedding in Steiner triple systems and Steiner quadruple systems respectively can also be treated as partial cases of this theorem.

6 Acknowledgements

This research was financed by the Russian Science Foundation (grant No 14-11-00555).

References

- [1] S. V. Avgustinovich and D. S. Krotov. Embedding in a perfect code. *J. Comb. Des.*, 17(5):419–423, 2009. DOI: 10.1002/jcd.20207.
- [2] B. Ganter. Finite partial quadruple systems can be finitely embedded. *Discrete Math.*, 10(2):397–400, 1974. DOI: 10.1016/0012-365X(74)90130-7.
- [3] K. T. Phelps and M. Villanueva. Ranks of q -ary 1-perfect codes. *Des. Codes Cryptography*, 27(1-2):139–144, 2002. DOI: 10.1023/A:1016510804974.
- [4] A. M. Romanov. On the admissible families of components of Hamming codes. *Diskretn. Anal. Issled. Oper.*, 19(2):84–91, 2012. In Russian.
- [5] C. Treash. The completion of finite incomplete Steiner triple systems with applications to loop theory. *J. Comb. Theory, Ser. A*, 10(3):259–265, 1971. DOI: 10.1016/0097-3165(71)90030-6.